

Ontology for Level-One Sensor Fusion and Knowledge Discovery

Marion G. Ceruti, Ph.D.

Space and Naval Warfare Systems Center, San Diego (SSC-SD), Code 24121
53560 Hull Street, San Diego, CA 92152-5001
canzone268@earthlink.net, marion.ceruti@navy.mil

Abstract. This position paper presents concepts and applications of an ontology for level-one sensor fusion. It covers platforms, sensors, tangible entities as well as intangible entities, such as concepts. Relationships between entities are part of the ontology. Verbs as well as nouns are included. Concepts such as reasoning under uncertainty also are part of the sensor-fusion ontology. Applications relate to knowledge discovery and the pattern recognition of potential security threats.

1 Introduction

Sensor-data fusion is divided into four levels of increasing situation complexity. Level-one sensor data fusion, which concerns object refinement, is defined as the fusion of data related to detection, tracking, classification and the identification of platforms such as ships and aircraft without consideration to the intent of the platforms. (See, for example, [1]). Level two is focused on situation refinement in which the relationship between platforms becomes important. Level three is about threat refinement and addresses the intent of hostile platforms. Level four addresses process refinement in which a commander tries to predict hostile actions. Knowledge discovery and the fusion of sensor data with information from a variety of other observations can assist military and law-enforcement efforts to detect anomalous behavior of platforms, thus contributing to security and threat detection at seaports and airports.

A comprehensive ontology for level-one sensor fusion includes several sub-levels of ontology and several different dimensions of level-one fusion: platforms and sensors, characteristics, tangible and intangible, nouns and verbs, relationships between variables, concepts such as data combinations. For example, the speed of a ship and its position are to a first approximation, independent data. Knowing only the speed of a ship may not trigger an alert. Similarly, its position alone may not be significant. However, the knowledge-discovery process, may reveal that the speed of a particular ship located in a certain area is unusual could signify illegal activity.

Data sources on a sensor ontology include data dictionaries that contain terms associated with multiple sensor types such as acoustic, magnetic, visual, imagery, electro-optical etc., as well as other sources that describe how sensors work alone and together, such as is the case with sensor fusion. This paper defines concepts and specifies relationships between entities for the objects and concepts.

The military and law-enforcement organizations need a single, integrated, logical and national sensor ontology to support knowledge bases in expert systems designed for joint use and for homeland security. It represents the future for fusion, sensor networks and intelligence. Existing ontologies, like databases, are fragmented, incomplete and in different formats. This work was performed as part of a project at SSC-SD to test and evaluate Building the Single Integrated Picture (BSIP) [2].

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 24 SEP 2004		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Ontology for Level-One Sensor Fusion and Knowledge Discovery Proceedings of the 2004 International Knowledge Discovery and Ontology Workshop (KDO-2004), pp. 97-102, 20-24 Sep. 2004, Pisa Italy.				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Dr. Marion G. Ceruti				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) SSC Pacific San Diego, CA 92152-5001				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This position paper presents concepts and applications of an ontology for level-one sensor fusion. It covers platforms, sensors, tangible entities as well as intangible entities, such as concepts. Relationships between entities are part of the ontology. Verbs as well as nouns are included. Concepts such as reasoning under uncertainty also are part of the sensor-fusion ontology. Applications relate to knowledge discovery and the pattern recognition of potential security threats.					
15. SUBJECT TERMS data fusion, data sources, platforms, sensor-data fusion, sensor ontology, signals, tracking					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 6	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

2 Symbology and levels of Abstraction

Ontologies themselves are an abstraction from physical objects and actions. Ontologies handle multiple levels of abstraction in a hierarchical manner. Sensor fusion has a particular challenge in that it involves a meta-abstraction. For example, data about platforms and sensors are an abstraction from physical platforms and sensors. The data are represented as different icons in decision-support systems for different user communities. This second level of abstraction is a meta-abstraction.

Every level of abstraction introduces additional error modes, especially with level-one fusion. To reduce difficulty and confusion, information exchange needs to take place at the same level of abstraction. For example, user communities should communicate data rather than specify the conventions used to represent data on displays. A symbol that is a yellow in one community is white in another.

The symbology of a community depends on the assumptions that the community makes about the data and platforms that the symbols are supposed to represent. Communities that specialize in force planning operate in an analytical mode and prefer complex symbols that are designed to communicate relatively large quantities of information at a glance to personnel who know how to interpret them.

In contrast, in a user community that deals with tactical operations in which decisions need to be made in time-critical scenarios, the symbols tend to be simpler. For example, a hostile submarine is represented simply as a red, inverted “V” with a dot in the middle of it to show the assumed position. This symbol system is designed to alert an operator to danger, not to facilitate long-term planning [3].

An ontology that supports level-one sensor fusion includes concepts from the planning and operational communities. Different groups conceptualize and use sensor data in different ways. Knowledge discovery for detecting complex patterns in the battle space or during peace-time security missions will need to account for the different levels of granularity and detail that are used in planning and operations.

3 Relationships between Entities

Nouns represent the entities and verbs contribute to the relationships between them. For example, some important relationships include basic verbs such as “is a,” “is a kind of,” “has a,” “is between,” “originates from” and “detects.” In addition to verbal relationships, other interactions such as the speed-direction interaction mentioned above, and the speed-location interaction also can be important in a knowledge-discovery task. In general, a complete ontology also supports concepts based on inverse relationships, such as “If A sends C to B, C must have been at one time between A and B.” In addition to the general verbs used in any ontology, certain verbs apply especially to sensors. The verb, “detect,” is quite basic to sensors. Other verbs are “identify,” “classify,” “occlude,” “attenuate,” “amplify” and “fuse.”

Semantic relationships between actions play an important and often key role in sensor-related concepts. Finer grained relationships are also possible, such as strength, enablement and temporal information [4]. For example, X may happen before Y. X could be detect and Y could be fuse. In this case X enables Y because Y cannot occur without first having X. It also implies that X must occur BEFORE Y, thus invoking the concept of temporal information being stored in verb usage. Any

comprehensive knowledge discovery engine must process relationships between entities and the verbs that are part of and that influence these relationships.

An example of strength relations in sensor verbs, is the fact that “identify” is stronger than “classify,” which in turn is stronger than “detect.”

“Happens before” is another semantic relation [4]. For example, when referring to a sensor-data-acquisition task, “acquire” happens before “store.” Also, platform “detection” happens before “classification” or “identification.” Thus, “detect,” “classify,” and “identify,” are related not only by verb strength but also by the “happens before” semantic relation. These concepts are described in section 7.

Antonymy [4] also is present in sensor concepts. For example, “amplify” is the opposite of “attenuate” when used in reference to signals.

In addition to verbs, many more nouns apply to an ontology. For example the noun ontology includes the following terms and the relationships between them: geometry, detector, platform, range, signal, sensor, mode, fusion, uncertainty, and error. The next sections cover some of these concepts in more detail.

4 Platforms – Concepts and Hierarchy

The general category of platforms can be divided along several dimensions. Among them are: stationary or capable of motion; machine powered and wind powered; aircraft, vessels, land vehicles and amphibious vehicles; modern and archaic; driven by humans vs. autonomous. The main feature that all platforms have in common in the context of target detection is that all platforms are capable of motion or at least they were capable of motion in the recent past. Moreover, a platform must be capable of transporting another entity, such as humans, a payload, cargo, or any combination of these. This definition automatically excludes ships and aircraft that have been brought onto land in a display status and are no longer functional for transportation and, therefore are of little to no interest in terms of target detection.

The characteristics and performance of platforms constitute the main source of information from which platforms are detected, classified and identified. Thus, a platform ontology will have two main branches, one for nouns (what it is and has) and the other for verbs (what it does). For target detection, an essential feature of platforms is that they can be divided into parts, both conceptually and physically. The meronymy of a platform for signal detection purposes includes those particular parts that give rise to signals that can reveal the existence of, classify, or identify the platform. The “partology” of a platform is an important part of signal-detection ontology.

Performance is a good source of information on platforms, but performance alone often cannot provide enough information to complete a target-identity determination. Information about a platform can be revealed by its velocity, which consists of two parts, speed and direction. The speed of a platform can reveal the class to which it belongs by excluding any platform with a maximum speed less than the observed speed of the platform. The direction of motion also can reveal certain information about the platform. The observed speed and location of platform along with its notional capabilities (e.g. top speed, operating depth, minimum depth required for draft, etc.) together may provide enough information to classify a platform.

5 Signals

For level-one fusion, signals are emissions from platforms that have a certain form of energy, such as acoustic, magnetic, electro-optic, radar, etc. This energy is characterized by several features, such as peak frequency, if pulsed, pulse repetition rate, speed (which varies with signal type and medium), spectral characteristics, onset, and duration. Signals can result from passive detection, as is the case with magnetic signatures, or from active sonar or radar, in which another entity, usually another platform supplies the initial source of energy and the reflection of this emission which is altered by the platform of interest is analyzed. The assumption on which tasks of detection, classification and identification are based is that by analyzing signals, including visual signals analyzed directly by the human nervous system, one can determine required information (within a particular timeframe) about the platform that was the source or reflection of these signals. In many cases this is a true assumption. In others, too much uncertainty prevents a timely determination.

The following concepts are a part of the signal ontology: emission, reflection, energy type (e.g. acoustic or electromagnetic), energy characteristics like spectral features (e.g. frequency, pulse repetition rate, inter-pulse interval, amplitude, wave speed and wavelength), imagery and its patterns and features, and reasoning under uncertainty. Other important concepts to be included in a signal ontology are signal jamming, signal processing, noise, signal-to-noise ratio, and detection.

6 Sensors

A *sensor* as a device that responds to a physical stimulus, (such as heat, light, sound, pressure, magnetism, or a particular motion) and transmits a resulting impulse for measurement or operating a control. In level-one fusions, sensors detect signals from platforms. Sensors can be mechanical, electrical, electro-optical, chemical, or biological (part of a living organism). Sensors can be land based or platform based, such as towed arrays. The ontology includes concepts such as signals, detection, modes of operation, detection threshold, tasking of sensors, sensor arrays, operational limitations (e.g. range, depth, distance and altitude), swath, resolution, passive sensors (e.g. hydrophone array), and active sensors (e.g. sonar). Concepts that deal with the interaction of signals with sensors include detection and signal-to-noise ratio mentioned in the section on signals. Thus, depending on the ontological representation either multiple paths will link these concepts or they will occur twice in the ontology.

Sensor geometry is also important, as some sensors are not omni-directional. Their geographical environment limits others. Thus concepts that apply to geometry in general, such as points, lines, line segments, polygons, circles, ellipses, and other spatial entities are also a part of the ontology that applies to sensors in particular. Some examples of spatial relationships and geometry as applied to sensors are as follows: 1. An ellipse is drawn around a position to indicate the error in the position measurement. The major and minor axes of the ellipse represent the uncertainty in the coordinates of the position. 2. If a sensor must detect a target, the target must be within the range of the sensor, both in the distance (radius) and angularly (azimuth). Three-dimensional geometry is important when a sensor is located in a canyon and it depends upon a line-of-sight geometry to detect a target.

7 Data Fusion and Tracking

The branch of the ontology that deals with data fusion and tracking includes the following concepts: sensor data fusion, fusion algorithms, track fragmentation, data perishability detection vs. classification, and identification vs. classification. Data fusion includes the concept of correlating multiple observations of the same platform with different sensors, or with the same sensor at different times along the platform's trajectory. This is known as tracking. When something interrupts the tracking process this may result in track fragmentation. There is no guarantee that all track fragments that originate from the observation of a given platform will be identified as having an association with that platform.

The notion of data perishability invokes the concept of time sensitivity. Time criticality is discussed in [5]. This concept relates to the idea that data that are useful now may not be of any utility in the near future, when the situation can change rapidly as platforms move in and out of range of sensors. Time-sensitive data are useful in an operational sense only as long as a tactical action can be performed in which the knowledge of these data plays a significant role. After a certain time period has expired, these data may remain useful for the purposes of archives and strategic analysis, but the operational picture has changed significantly that would render these data irrelevant for operational applications.

Detection is the act of becoming aware of the presence of a platform or a target is known. Classification occurs when the platform is put into a known category. Identification occurs when the platform can be named, or its serial number becomes known and all other platforms can be excluded, even those in the same class. These concepts occur in the ontology as a sequence. Identification and classification include detection, as detection must occur before a platform can be classified.

Tracking involves uncertainty, which is part of the sensor ontology covers the following concepts: epistemological concepts, percent probability, hypothesis formation, hypothetical platform, proposed ship identity, detection vs. classification, identification vs. classification, confirming evidence, and error. For example, do I see one platform as several tracks with disjoint multiple parts from various sensors, or do these disjoint tracks represent more than one target?

8 Knowledge Discovery in Sensor Data – Future Research

A sensor ontology can contribute value to the knowledge-discovery process in sensor data. Vast quantities of data that were collected over the years from sensors during battles and exercises have fallen by the wayside because no one has had the tools, the resources, or the inclination to interpret their significance. Without better data organization and advanced tools, such as those based on artificial intelligence, the problem has been too hard to solve. Patterns hidden in the data could be revealed much more efficiently if these data were organized into knowledge based on a sensor ontology. Thus, the data could be subject to artificial intelligence tools for knowledge discovery that otherwise could not be used due to various constraints [5, 6].

For example, knowledge-discovery tasks include association, sequential patterns, classification, and clustering [6]. Knowledge of the sensor ontology is a critical phase

of the knowledge-discovery process because it will prevent associating entities that are semantically distant but lexically close or identical. For example, the geometry of a sensor-platform configuration may be measured in nautical miles, abbreviated NM. The wavelength of electromagnetic radiation measured from a platform can be expressed in nanometers, also abbreviated NM. The association of a geometry with a wavelength only on the basis that they are both abbreviated NM is an erroneous result. However in the absence of a sensor ontology that would clearly separate the two concepts, some knowledge discovery software may be inclined make the association.

A sequential-pattern detection data-mining task is particularly well suited to the reconstruction of tracks from a database in multi-sensor correlation and tracking. This technique can be used to analyze the correctness of past track identifications that may have been the result of a few readily available techniques in operational scenarios. With more powerful knowledge-based techniques, pattern detection can be revisited on data sets from past exercises, etc. The knowledge base will be based on the sensor ontology. Pedigree data also are important here. Since records from different sources need to be related, these records need to be tagged with the identity of the entity that they represent [6]. As stated above, classification is not only a level-one sensor-fusion task, but also a data-mining task. Platform classifications that were missed in the operational environment due to time and other resource constraints can be discovered in an analytical mode after the battle or exercise during a data-mining classification task aimed at evaluating the efficiency of the operational data classification tools. A sensor ontology can help to validate similarity and differences of entities.

Acknowledgements

The author thanks the Office of Naval Research and the Space and Naval Warfare Systems Center, San Diego, Science and Technology Initiative for their support of this work. This paper is the work of a U.S. Government employee performed in the course of employment and no copyright subsists therein.

References

- [1] E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Boston 1990.
- [2] M.G. Ceruti and J.L. Kaina, "Enhancing Dependability of the Battlefield Single Integrated Picture through Metrics for Modeling and Simulation of Time-Critical Scenarios," *Proc. of the IEEE 9th Intl. Workshop on Real-time Dependable Systems, (WORDS 2003F)*, Oct. 2003.
- [3] Lead Standardization Activity (LSA) Center for Standards (CFS), *DOD Interface Standard Common Warfighting Symbolology*, MIL-STD-2525B, Reston, VA, 30 Jan. 1999.
- [4] T. Chklovski and P. Pantel, "Large-Scale Extraction of Fine-Grained Semantic Relations Between Verbs," 2004.
- [5] M.G. Ceruti and M.N. Kamel, "Preprocessing and Integration of Data from Multiple Sources for Knowledge Discovery," *International Journal on Artificial Intelligence Tools*, vol. 8, no. 3, pp. 159-177, June 1999.
- [6] M.N. Kamel and M.G. Ceruti, "Knowledge Discovery in Heterogeneous Environments," Chapter 13 in *Heterogeneous Information Exchange and Organizational Hubs*, ed. H. Bestougeff, et al., pp. 181-204, Kluwer Publishers, New York, June 2002.